

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 5, 2003, 01:50:12 ; Search time 287 seconds
(without alignments)
3452.541 Million cell updates/sec

Title: US-09-847-081B-2

Perfect score: 2270
Sequence: 1 MSMSVALLWVSPTEVSG.....IYAKSLVPPNRTSSPLAKT 440

Scoring table: BLOSUM62 Xgapop 10.0, Xgapext 0.5
Xgapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-Q/cgn2_1/USPTO.spool/US09847081/runat_01042003.120128.26729/app_query.fasta_1.583
-DB-N Geneseq 101002 -OPMT-fastap -SUFFIX-rng -MINARCH-0.1 -LOOPCL-0
-LOOPEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-blosum62 -TRANS-human40.cdi
-LIST-45 -DOCALLIGN-100 -THR_SCORE-pct -THR_MAX-100 -THR_MIN-0 -ALIGN-15
-MODE-LOCAL -OUTFMT-pto -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-2000000000
-USER-US09847081.ecgn_1.1.396_runat_01042003.120128.26729 -NCPU-6 -ICPU-3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES-0 -WAIT -LONGLOG -DEV TIMEOUT-120
-WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6 -FGAPEXT-7
-YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database : N_Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2270	100.0	1728	24	AAI66366
2	1826	80.4	1826	19	AAV03880
3	1801	79.3	1712	24	AAI66367
4	1799	79.3	1814	19	AAV03881
5	1774	78.1	1795	19	AAV03878
6	1762	77.6	1316	19	AAV03879
7	1751	77.1	1591	17	AAQ9323
8	1744	76.8	1239	19	AAV17247
9	1737	76.5	1239	21	AAZ99482
10	1729	76.2	1646	12	AAQ12495
11	1595	70.3	2868	19	AAV16951
12	1557	68.6	1921	19	AAV16949
13	1542	67.9	1566	21	AAAC48162
14	1542	67.9	1703	21	AAAC35120
15	1531	67.4	2085	19	AAV16948
16	1523	67.1	1932	19	AAV16950
17	1514	66.7	2585	23	ABLA1600
18	1431.5	63.1	1304	24	ABA97361
19	1393	61.4	1397	21	AAZ29145
20	1252.5	55.2	1448	21	AAZ29139
21	1078	47.5	1021	21	AAZ29146
22	1033.5	45.5	3485	13	AAQ29121
23	1033	45.5	992	21	AAZ29144
24	992	43.7	1060	21	AAZ29143
25	781.5	34.4	888	21	AAZ29140
26	511	22.5	476	21	AAZ29142
27	468	20.6	684	22	AAH44248
28	458.5	20.2	1509	15	AAQ64910
29	442	19.5	749	19	AAV03882
30	420	18.5	766	21	AAZ29141
31	378	16.7	56609	21	AA81459
32	378	16.7	349980	21	AAF21609
33	378	16.7	1437668	21	AAA81490
34	364	16.0	1198	12	AAQ13718
35	364	16.0	1198	17	AAAT40791
36	364	16.0	1198	17	AAAT41743
37	364	16.0	1198	17	AAAT37093
38	364	16.0	1198	18	AAAT91543
39	363.5	16.0	930	18	AAAT66534
40	350.5	15.4	1232	19	AAV19122
41	350.5	15.4	1232	20	AAZ25063
42	350.5	15.4	1232	24	ABA97354
43	349.5	15.4	6918	11	AAQ06299
44	347.5	15.3	1386	13	AAQ29122
45	344.5	15.2	930	18	AAZ23885

ALIGNMENTS

RESULT 1	
AAI66366	
ID	AAI66366 standard; cDNA; 1728 BP.
XX	
XX	AAI66366;
XX	
XX	29-JAN-2002 (first entry)
DT	
DE	Nicotiana tabacum phytoene synthase coding sequence #1.
XX	
XX	Phytoene synthase; zeta carotene desaturase; herbicide; transgenic plant;
KW	plant growth regulator; herbicidal; tobacco; ss.
XX	
OS	Nicotiana tabacum.
XX	
Key	Location/Qualifiers
FT	244...1566
CDS	/*tag- a
FT	

/product= "phytoene synthase"

FT XX DE10022362-A1.
 PN XX 15-NOV-2001.
 PD XX
 XX 08-MAY-2000; 2000DE-1022362.
 PF XX
 XX 08-MAY-2000; 2000DE-1022362.
 PR XX
 XX (FARB) BAYER AG.
 PA XX
 XX Busch M, Hain R;
 PI XX
 XX WPI: 2002-027336/04.
 DR XX P-PSDB; AM51841.
 DR XX
 XX
 PT New nucleic acid encoding tobacco zeta-carotene desaturase, useful for
 PT screening compounds with herbicidal activity -
 PT
 PS Claim 14; Page 12-17; 44pp; German.
 PS
 XX The present invention provides the protein and coding sequences of
 CC phytoene synthases and zeta-carotene desaturase from *Nicotiana tabacum*.
 CC The sequences can be used to identify compounds capable of altering the
 CC expression of these genes, which are therefore useful as plant growth
 CC regulators and herbicides. They can also be used to produce transgenic
 CC plants. The present sequence is the coding sequence of a tobacco
 CC phytoene synthase.
 CC
 XX
 SQ Sequence 1728 BP; 513 A; 290 C; 433 G; 492 T; 0 other;
 Alignment Scores:
 Pred. No.: 7,898-234 Length: 1728
 Score: 2270.00 Matches: 440
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 US-09-847-081B-2 (1-440) x AA166366 (1-1728)
 QY 1 MetSerMetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGly 20
 DB 244 ATGAGCATGCTGCTTGTGTTGGTGTGTTCTCCACATCCGAGGCTCGAATGG 303
 QY 21 ThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu 40
 DB 304 ACAGGATTGTTGGATTTCAGTCCGAGAAAGAAACCGCGCTTTGTATCATCCAGGTTCTTA 363
 QY 41 AlaArgAspArgAsnLeuMetTrpAsnGlyArgGlyLeuLysGlyGlyArgGlnArgTrp 60
 DB 364 GCTCGAGATAGGAATTTGATGTGGATGGAGATCAAGAAAGGTGGGAGACAAAGGTGG 423
 QY 61 AsnPhcGlySerLeuLeuAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGlu 80
 DB 424 AATTTTGGCTCTTAACTGCTGATCCAGATATTCATGCTTGGTGGATCAAGAACTGAA 483
 QY 81 LysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThr 100
 DB 484 AAGGAAGACATTTCTCTGTACAGTCCAGTTGGTGGCTAGCCAGCTGGAGAAATGACT 543
 QY 101 ValSerSerGluLysLysValTrpAspValValLeuLysGlnAlaAlaLeuValLysArg 120
 DB 544 GTGTATCAGAGAAAGGTGATGATGTTGATTAAGCAGGCGCTTTAGTAGAGAGG 603
 QY 121 GlnLeuArgSerThrAspAspLeuGluValLysProAspIleValValProGlyAsnLeu 140
 DB 604 CAGCTCAGATCTACCGATGATTTAGAAAGTGAAGCCGATGATTGTTCTCCAGGAATTG 663
 QY 141 GlyLeuLeuSerClnAlaLysArgCysGlyGluValCysAlaGluTrpAlaLysThr 160
 DB 664 GGCTTTGAGTGAGCATATGATCGTTGTGGCAAGATGTGTCAGAGATGTCAAAGACA 723

QY 161 PheTyrLeuGlyThrLysLeuMetThrProGluArgArgAlaIleTrpAlaIleTyr 180
 DB 724 TTTTACTTAGGAACCAAGCTAATGACCCAGAGAGAGAGAGCTATCTGGGCAATATAT 783
 QY 181 ValTrpCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrPro 200
 DB 784 GTGTGGTGGAGAGACGGATGAGCTTGTGTGGCCCTAATGCATCCCACTAATCTCCG 843
 QY 201 GlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPheAsp 220
 DB 844 CAAGCTTTAGATAGGTGGGAGACAGGCTGGAGATATTTTCACTGGCGGCCATTGAT 903
 QY 221 MetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPhe 240
 DB 904 ATGCTTGATGCTGCTTTATCCGATACCTCTCCAGATTTCTCTGTGATATTCAGCCATTC 963
 QY 241 ArgAspMetIleGluGlyMetArgMetAspLeuTyrLysSerArgTyrLysThrPheAsp 260
 DB 964 AGAGATATGATTGAGGAATGGTATGGACTTGTGGAAATCCAGATACAAAACCTTCGAT 1023
 QY 261 GluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProVal 280
 DB 1024 GAGCTATATCTCTATTGTTACTATGTTGCTGTTACTGTTAGGATTTGATGCTTCCAGTT 1083
 QY 281 MetGlyIleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAla 300
 DB 1084 ATGGGTATTGCACTTCAATCAAGGCAACAGAGAGTGTATATATGCTGCTTTGGCT 1143
 QY 301 LeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArg 320
 DB 1144 TTAGGCTTGCATCAATCACTAATCAATATATCTCAGAGATGTAGGAGAAAGATGCCAGAGA 1203
 QY 321 GlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIle 340
 DB 1204 GGAAGAGTATATCTGCTCAAGATGAATTAGCACAGCGAGGCTCTCCGACGAAGACATA 1263
 QY 341 PheAlaGlyArgValThrAspLysTyrPheAsnPhenMetLysGlnIleGlnArgAla 360
 DB 1264 TTTGCTGGAGAGTGTACTGATAGTGGAGAACTTTATGAAGAAACAAATTCAGAGGGCG 1323
 QY 361 ArgLysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrp 380
 DB 1324 AGSAAATCTTGTGATGAGTCAGAGAAAGGTGTACAGAACTGGACTCTGCTAGTAGTGG 1383
 QY 381 ProValLeuThrAlaLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAsp 400
 DB 1384 CCTGTGTTAACAGCGCTGCTGTGTATCGCAAGATATTGGACGAGATTGAAGCCACAGC 1443
 QY 401 TyrAsnAsnPhenThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeuPro 420
 DB 1444 TACACAACTTCACAGGAGGCTTATGTAGCAAGCCAAAGAGCTTCTCACCCTTGCCC 1503
 QY 421 IleAlaTyrAlaLysSerLeuValProProAsnArgThrSerSerProLeuAlaLysThr 440
 DB 1504 ATTGCTTATGCAAAATCTCTTGTGCCCCCTAATAGAATCTCCTCTCCACTAGCAAGACA 1563

RESULT 2

AAV03880

ID AAV03880 standard; cDNA; 1826 BP.

XX

AAV03880;

XX

XX 29-APR-1998 (first entry)

XX

DE Phytoene synthase coding sequence from *N. tabacum*.

XX

KW Phytoene synthase; transgenic plant; enhanced carotenoid synthesis;
 ultra violet absorber; food colour; ss.

XX

OS *Nicotiana tabacum*.

XX

FH Key Location/Qualifiers

FT CDS 367..1599
 XX /*tag= a
 PN US5705624-A.
 XX
 PD 06-JAN-1998
 XX
 XX 27-DEC-1995; 95US-0579667.
 XX
 XX 27-DEC-1995; 95US-0579667.
 XX
 PA (DELLA) DELLA-CIOPPA G R.
 PA (FITZ) FITZMAURICE W P.
 PA (GRILL) GRILL L K.
 PA (HELL) HELLMANN G M.
 PA (KUMAI) KUMAGAI M H.
 XX
 PI Della-Cioppa GR, Fitzmaurice WP, Grill LK, Hellmann GM;
 PI Kumagai MH;
 XX
 DR WPI; 1998-086196/08.
 DR P-PSDB; AAW41059.
 XX
 PT DNA encoding tobacco phytoene synthase polypeptides - useful for
 PT producing recombinant polypeptides or transgenic plants
 XX
 PS Claim 1; Column 27-30; 25pp; English.
 CC This sequence encodes the phytoene synthetase from Nicotiana tabacum.
 CC The phytoene synthetase coding sequence represents a cDNA of the
 CC invention. The isolated nucleic acid molecules are used for producing
 CC recombinant polypeptides or transgenic plants with enhanced ability to
 CC synthesise carotenoids. Phytoene has been used as a ultra violet absorber
 CC and other carotenoids have been used as food colours, animal feeds and in
 CC the pharmaceutical and cosmetics industries.
 XX
 SQ Sequence 1826 BP; 589 A; 282 C; 433 G; 522 T; 0 other;
 Alignment Scores:
 Pred. No.: 4.61e-186 Length: 1826
 Score: 1826.00 Matches: 362
 Percent Similarity: 88.61% Conservative: 27
 Best Local Similarity: 82.46% Mismatches: 21
 Query Match: 80.44% Indels: 29
 DB: 19 Gaps: 4
 US-09-847-081B-2 (1-440) x AAV03880 (1-1826)
 QY 3 MetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGlyThrGly 22
 DB 367 ATGTCGTGCTGCTTGTATGGTGTGTTTCACT---TGTGAGGCTCAATGGGACAGA 423
 QY 23 LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42
 DB 424 TCTTGGATTTCAGTAAGGGAGGAAACCGGGTTTGTGCTGCTG-----AGG 471
 QY 43 AspArgAsnLeuMetTrpAsnGlyArgIleLysGlyGlyArgGlnArgTrpAsnPhe 62
 DB 472 CATGAGATTAGTGTGCAATGAGAGATCAAGAGAGGTGTGAACAAAGGTGGATTTT 531
 QY 63 GlySerLeuIleAlaAspProArgTrpSerCysLeuGlyGlySerArgThrGluLysGly 82
 DB 532 GGT----- 534
 QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMet---ThrVal 101
 DB 535 -----TCTGTACGCTCTCGCATGTGTGTACACCAACGGGAGAAATGGCGACAATG 585
 QY 102 SerSerGluLysValTrpAspValValLeuLysGlnAlaLeuValLysArgGln 121
 DB 586 ACATCAGACAGAGGTTTATGATGTGTTGTAACAGCAGCTTTAGTGAAGAGGAG 645
 QY 122 LeuArgSerThrAspAspLeuGluValLysProAspIleValValProGlyAsnLeuGly 141

Db 646 CTGAGATCTACTGATGATTTAGAAAGTGAAGCGGAGATCCCTCTCCCGGAATTTGAGC 705
 QY 142 LeuLeuSerGluAlaTrpAspArgCysGlyGluValCysAlaGluTrpAlaLysThrPhe 161
 Db 706 TTGTTAAGTGAAGCATATGATAGGTGTAGTGAAGTATCGCAGATATGCAAGACATTT 765
 QY 162 TyrLeuGlyThrLysLeuMetThrProGluArgArgAlaIleTrpAlaIleTyrVal 181
 Db 766 TACTTAGAACTATGCTAATGACTCCACAGAGAAGCGCTATTGGCAATATATGTA 825
 QY 182 TrpCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGln 201
 Db 826 TGGTGCAGGAGACAGATGAACCTTGTGATGCCCGGATGTCATCATATTAATCAACA 885
 QY 202 AlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMet 221
 Db 886 GCCTTAGATAGTGGGAAGACCGGCTGGAAGATGTTTTCAGTGGCGGCGCATTTGATATG 945
 QY 222 LeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArg 241
 Db 946 CTCGATGCTGCTTGTCCGATACCTGTTCCAGTTTCCAGTTGATATTCAGCGGTCAGA 1005
 QY 242 AspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyrLysThrPheAspGlu 261
 Db 1006 GATATGATTGAAGAAATCGTATGGACTTGAGGAAGTCAAGATACAGAACTTTGATGAG 1065
 QY 262 LeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMet 281
 Db 1066 CTATACCTATATTGTTTATGCTGTTGCTGAGGTGGTGTGATGAGTGTTCCAATATG 1125
 QY 282 GlyIleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaLeuAlaLeu 301
 Db 1126 GGTATGCACTGATTCAGGCAACACAGAGAGTGTATATATGCAGCTTTGGCTTTA 1185
 QY 302 GlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgGly 321
 Db 1186 GGGATCGCAATCAACTAACCAACATACCTCAGAGATGTCGGAGAAGATCCAGAGAGCA 1245
 QY 322 ArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePhe 341
 Db 1246 AGAGTCTACTTACCTCAAGATGAGTTAGCACAGGAGGCTCTCTCCGACAAATGACATTTT 1305
 QY 342 AlaGlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGlnArgAlaArg 361
 Db 1306 CTGGAAGTGTGCTGATTAATGGAGAAGCTTTATGAAGAACCAATCCAGAGGGCAAGA 1365
 QY 362 LysPheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrpPro 381
 Db 1366 AAATTTCTCGAGGAGGAGGAGGAGGAGTGCACAACTCAGCTCAGCTAGTAGATGGCCT 1425
 QY 382 ValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAspTyr 401
 Db 1426 GTATGGGCATCTTGTCTGTGTACCGCAGATACCTCGACGAGATTGAAGCCATGACTAC 1485
 QY 402 AsnAsnPheThrArgArgAlaTrpValSerLysProLysLysLeuLeuThrLeuProfile 421
 Db 1486 AACAACTTCACAGGAGAGCTTATGTGAGCAAAACCAAGAACTTAATTTCTTACCTATT 1545
 QY 422 AlaTyrAlaLysSerLeuValProAsnArgThr-SerSerProLeuAlaLys 439
 Db 1546 GCTTATGCAAAATCTTTGTGCCCCCTACAAGAACTCTTGTTCACCTCTAGCTAAG 1600
 RESULT 3
 AAI66367
 ID AAI66367 standard; cDNA; 1712 BP.
 XX
 AC AAI66367;
 XX
 DT 29-JAN-2002 (first entry)
 XX
 DE Nicotiana tabacum phytoene synthase coding sequence #2.
 XX

Phytoene synthase; zeta: carotene desaturase; herbicide; transgenic plant;
plant growth regulator; herbicidal; tobacco; ss.

Nicotiana tabacum.

Key Location/Qualifiers
CDS 333..1565
/*tag-a
/product= "phytoene synthase"

DE10022362-A1.

15-NOV-2001.

08-MAY-2000; 2000DE-1022362.

08-MAY-2000; 2000DE-1022362.

(FARB) BAYER AG.

Busch M, Hain R;

WPI; 2002-027336/04.

P-PSDB; AAM51842.

New nucleic acid encoding tobacco zeta-carotene desaturase, useful for
screening compounds with herbicidal activity -

Claim 14; Page 21-26; 44pp; German.

The present invention provides the protein and coding sequences of
phytoene synthases and zeta-carotene desaturase from *Nicotiana tabacum*.
The sequences can be used to identify compounds capable of altering the
expression of these genes, which are therefore useful as plant growth
regulators and herbicides. They can also be used to produce transgenic
plants. The present sequence is the coding sequence of a tobacco
phytoene synthase.

Sequence 1712 BP; 540 A; 265 C; 421 G; 475 T; 11 other;

Alignment Scores:

Pred. No.:	2,046-183	Length:	1712
Score:	1801.00	Matches:	357
Percent Similarity:	87.93%	Conservative:	29
Best Local Similarity:	81.34%	Mismatches:	24
Query Match:	79.34%	Indels:	29
DB:	24	Gaps:	4

US-09-847-081B-2 (1-440) x AAI66367 (1-1712)

QY	3	MetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGlyThrGly	22
DB	333	ATGCTGTTGCTTGTATGGGTGTTTACCT--TGTAAGTCTCAATGGGACAGA	389
QY	23	LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg	42
DB	390	TTCTTGGAATTCAGTCGGGAGAACCGGTTTGTATGCTCG	437
QY	43	AspArgAsnLeuMetTrpAsnGlyArgLleLysLysGlyArgGlnArgTrpAsn	62
DB	438	CATAGGAATTTAGTGTCAATGAGAACACAGAGAGGTGTGAACAAAGGTGGAATTT	497
QY	63	GlySerLeuIleAlaAppProArgTyrSerCysLeuGlyGlySerArgThrGluLysGly	82
DB	498	GGT-----	500
QY	83	SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMet--ThrVal	101
DB	501	-----TCGTAGGCTCTGCTATGGTGGCTACACCGCGGGAGAAATGGGACGATG	551
QY	102	SerSerGluLysLysValTrpAspValValLeuLysGlnAlaLeuValLysArgGln	121
DB	552	ACATCAGACAGATGTTTATGATGGTGTGTTTAAACACAGCAGCTTTAGTGAAGGCGAG	611

QY	122	LeuArgSerThrAspAspLeuGluValValProAspIleValValProGlyAsnLeuGly	141
DB	612	TTGAGATCTGCTGATGATTAGAGTGAAGCGGAGATCCTCTCCCGGAATTTGAGC	671
QY	142	LeuLeuSerGluAlaTrpAspArgCysGlyGluValCysAlaGluTrpAlaLysThrPhe	161
DB	672	TTGTTGAGTGAAGCATATGATAGGTAGTGAAGTATGTGCAGAGTATGCAAGACATTT	731
QY	162	TyrLeuGlyThrLysLeuMetThrProGluArgArgAlaIleTrpAlaIleTyrVal	181
DB	732	TACTTHGGACCATGYTAATGACTCCAGAGAGAAAGGCTATTGGGCAATATATGTG	791
QY	182	TrpCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGln	201
DB	792	TGGTCAGGAGAACAGATGAACCTTGTATGGCCCAACCATCATATATACACCCCAA	851
QY	202	AlaLeuAspArgTrpGluThrArgLeuAspIlePheSerGlyArgProPheAspMet	221
DB	852	GCTTAGATAGGTGGGAGACCGGCTTGAAGATGTTTTCAGCGGGGACCATTTGATATG	911
QY	222	LeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArg	241
DB	912	CTCGATGCTGCTTGTCCGATACCTGTTTCCAACTTCCAGTTGATATTCAGCCGTTTCA	971
QY	242	AspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyrLysThrPheAspGlu	261
DB	972	GATATGATTGAAGGAATGCGTATGAGTGTGAGGAAGTCAAGATATAGAACTTTGATGAG	1031
QY	262	LeuTyrLeuTyrCysTyrTrpValAlaGlyThrValGlyLeuMetSerValProValMet	281
DB	1032	CITTACCTCTATTGTTTATTACGTGCTGTTGTTGGTGGTGGTGGTGGTGGTGGTGGT	1091
QY	282	GlyIleAlaProGluSerLysAlaThrThrGluSerValTrpAsnAlaAlaLeuAlaLeu	301
DB	1092	GGTATTGCACTGATTCAAAGGCAACACAGACAGAGCGGTATATATGACGCTTTGGCTTA	1151
QY	302	GlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGly	321
DB	1152	GGAAATCGCAATCAACTAACGAACATCTACAGAGATGTTGGGAAGATGCCAGAAGAGA	1211
QY	322	ArgValTrpLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePhe	341
DB	1212	ACAGTCTACTTACCTCAAGATGAATTTAGCACAGCGAGGCTCTCTCGACGATGACATATT	1271
QY	342	AlaGlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGlnArgAlaArg	361
DB	1272	GCTGGAAGAGTGACTGATAAGTGGAGAAAGCTTTATGAAGAAGCAAAATCCAGAGGCA	1331
QY	362	LysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrpPro	381
DB	1332	AASTTCTTCGATGAGGACAGGAGGAGGTATACAACTAGCTAGCTAGCTAGCTAGCT	1391
QY	382	ValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAspTyr	401
DB	1392	GTATGGGCATCTTGTGTTGTACCCCAATACTGACGAGATGAAGCAATGACTAC	1451
QY	402	AsnAsnPheThrArgArgAlaTrpValSerLysProLysLysLeuLeuThrLeuProIle	421
DB	1452	AACAACCTTCACAAAGAGAGCTTATGTGCAACCAACCAAGAAAGCTAATTTCTTACCTATT	1511
QY	422	AlaTyrAlaLysSerLeuValProProAsnArgThr-SerSerProLeuAlaLys	439
DB	1512	GCTTTATGCAAAATCTCTTGTGCCCCCTACAAAGAAGCTTTGTCTACCTCTAGCTAAG	1566

RESULT 4

AAV03881	ID	AAV03881 standard; cDNA; 1814 BP.
XX	XX	
AAV03881;	AC	
XX	XX	
29-APR-1998	DT	(first entry)
XX	XX	

QY 420 ProileAlaTyAlaLysSerLeuValProProAsnArgThr-SerSerProLeuAlaLys 439
 DB 1535 CCTATTGCTAAATGCAAAATCTCTTGTGCCCCCAAGAAGACTCTTCTCTCTAGCTAA 1594
 QY 439 s 439
 DB 1595 G 1595

RESULT 6

AAV03879
 ID AAV03879 standard; cDNA; 1316 BP.

XX AAV03879;
 AC AAV03879;

DT 29-APR-1998 (first entry)

XX Phytoene synthase coding sequence from N. benthamiana.

XX Phytoene synthase; transgenic plant; enhanced carotenoid synthesis;
 KW ultra violet absorber; food colour; ss.

XX Nicotiana benthamiana.

EH Key Location/Qualifiers
 FT CDS 1..1242
 FT /*tag= a

XX US5705624-A.

XX 06-JAN-1998.

XX 27-DEC-1995; 95US-0579667.

XX 27-DEC-1995; 95US-0579667.

XX (DELL/) DELLA-CIOPPA G R.

PA (FITZ/) FITZMAURICE W P.

PA (GRILL/) GRILL L K.

PA (HELL/) HELLMANN G M.

PA (KUNAI/) KUNAGAI M H.

PI Della-Cioppa GR, Fitzmaurice WP, Grill LK, Hellmann GM;
 PI Kumagai MH;

XX WPI; 1998-086196/08.

DR P-PSDB; AAW41058.

XX DNA encoding tobacco phytoene synthase polypeptides - useful for
 PT producing recombinant polypeptides or transgenic plants

XX Claim 1; Column 21-26; 25pp; English.

XX This sequence encodes the phytoene synthetase from Nicotiana benthamiana.
 CC The phytoene synthetase coding sequence represents a cDNA of the
 CC invention. The isolated nucleic acid molecules are used for producing
 CC recombinant polypeptides or transgenic plants with enhanced ability to
 CC synthesize carotenoids. Phytoene has been used as a ultra violet absorber
 CC and other carotenoids have been used as food colours, animal feeds and in
 CC the pharmaceutical and cosmetics industries.

SQ Sequence 1316 BP; 403 A; 221 C; 342 G; 350 T; 0 other;

Alignment Scores:

Pred. No.: 2,18e-179 Length: 1316
 Score: 1762.00 Matches: 347
 Percent Similarity: 87.04% Conservative: 29
 Best Local Similarity: 80.32% Mismatches: 28
 Query Match: 77.62% Indels: 28
 DB: 19 Gaps: 4

US-09-847-081B-2 (1-440) x AAV03879 (1-1316)

QY 3 MetSerValAlaLeuLeuTrpValSerProThrSerGluValSerAsnGlyThrGly 22

DB 1 ATGTCTGTGCTGCTTGTATGGTTGTTTCCACCT---TGTGAGGTCTCAATGGGACAGGA 57
 QY 23 LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42
 DB 58 TTCCTGGATTCAATCCGGAGGGAACCGGGTTTTTGATTGGTGC-----AGG 105
 QY 43 AspArgAsnLeuMetTrpAsnGlyArgIleLysLysGlyGlyArgGlnArgTrpAsnPhe 62
 DB 106 CATAGGAATTTAGTGTGCAATCAGAGAAACAAGAGAGGTGTGGAACAATGTGGAATTTT 165
 QY 63 GlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGluLysGly 82
 DB 166 GGT----- 168
 QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMet---ThrVal 101
 DB 169 -----TCTGTAAAGTCTCTATGTTGGTGTACACCGCGGGAGAAATGGCGCAGATG 219
 QY 102 SerSerGluLysLysValTyrAspValValLeuLysGlnAlaAlaLeuValLysArgGln 121
 DB 220 ACATCAGAACACAGATGTTTATGATGTGTGTTGATTGAACAACAACAGCTTTAGTGAAGAGCAG 279
 QY 122 LeuArgSerThrAspAspLeuValLysProAspIleValValProGlyAsnLeuGly 141
 DB 280 TTGAGATCTACTGATGATTTAGAAAGTGAAGCGGAGATCCCTCCCGGGGAATTTGAGC 339
 QY 142 LeuLeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaLysThrPhe 161
 DB 340 TTGTTGAGTGAAGCATATATAGGTGTACGAGAGTATGTCAGAGTATGCGAAGACATTT 399
 QY 162 TyrLeuGlyThrLysLeuMetThrProGluArgArgAlaIleThrPalIleTyrVal 181
 DB 400 TACTTAGGAACATATGCTATGACTCCAGAGAGAAGAGGCTATTTGGCAATATATGTA 459
 QY 182 TrpCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGln 201
 DB 460 TGGTGCAGGAGAACACAGCAACTTGTGTATGTCGCCCAATGCATCATATATTACTCCCAA 519
 QY 202 AlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMet 221
 DB 520 GCCTTAGATAGGTGGGAAGACCGCGTGAAGATGTTTTTCAGTGGCGGCATTTGACATG 579
 QY 222 LeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArg 241
 DB 580 CTCGATGCTGCTTTGTCCGATACTGTTTCCAAGTTTCCAGTTGATATTCAGCGCTTCAGA 639
 QY 242 AspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyrLysThrPheAspGlu 261
 DB 640 GATATGATCGAAGGAATGCGTATGGACTTGAGAACTCGAGATACAGAACTTTTGATGAG 699
 QY 262 LeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMet 281
 DB 700 CTATACCTATATGTTTATGCTGTGTACAGTTGGTGTGATGAGTGTTCACATTAATG 759
 QY 282 GlyIleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeu 301
 DB 760 GGCATCGCACCTGAATCAAGAAGCAACAACAGAGAGTGATATATGACGCTTTGGCTTTC 819
 QY 302 GlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGly 321
 DB 820 GGTATCGCGAATCAACTAACCAACATTTCTCAGAGATGTCGGAGAAGATGCCAAGAGGA 879
 QY 322 ArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePhe 341
 DB 880 AGAGTCTACTTACTCTCAAGATGAATTAGCACAGCGAGGTCTCTCCGACGATGACATATT 939
 QY 342 AlaGlyArgValThrAspLysTyrArgAsnPheMetLysLysGlnIleGlnArgAlaArg 361
 DB 940 ACTGGAAAGTGCATGAATAATGGAGAAGCTTTATGAAGAAGCAATTCAGAGGGCAAGA 999
 QY 362 LysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrpPro 381

QY 361 ArgLysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArgTrp 380
 DB 1266 AGAAGTTCTTTGATGAGCGAGAAAGCGGTGACAGAAATTGAGCTCAGCTAGTAGATTG 1325
 QY 381 ProValLeuThrAlaLeuLeuLeuTyrArgLysLysLeuAspGluLeuGluAlaAsnAsp 400
 DB 1326 CCTGATGGGATCTTTGGTCTTGACCGCAAAATAGTAGATGAGATTGAAGCCAATGAC 1385
 QY 401 TyrAsnAsnPheThrArgAlaLysValSerLysProLysLysLeuLeuThrLeuPro 420
 DB 1386 TACAACAACCTTCAACAAGAGAGCATATGTGAGCAAAATCAAGAAAGATTGATGCTTACCT 1445
 QY 421 IleAlaTyrAlaLysSerLeuValProAsnArgThrSerSer 435
 DB 1446 ATTGCATATGCAAAATCTTTGTGCTCTCTACAAAAAAGTGCCTCT 1490

RESULT 8
 AAV17247
 ID AAV17247 standard; DNA; 1239 BP.
 XX
 AC AAV17247;
 XX
 DT 28-MAY-1998 (first entry)
 XX
 DE MTOM5, modified phytoene synthase gene.
 XX
 KW MTOM5; phytoene synthase; chloroplast targeting sequence; enzyme; tomato;
 KW protein expression enhancement; transgenic plant; carotenoid synthesis;
 KW lycopene; ds.
 XX
 OS Lycopersicon esculentum.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1239
 FT /*tag= a
 FT /*note= "contains an intron"
 FT 1..960
 FT /*tag= b
 FT /*number= 1
 FT 961..990
 FT /*tag= c
 FT /*number= 1
 FT 991..1236
 FT /*tag= d
 XX
 PN WO9746690-A1.
 XX
 PD 11-DEC-1997.
 XX
 PF 23-MAY-1997; 97WO-GB01414.
 XX
 PR 07-JUN-1996; 96GB-0011981.
 XX
 PA (ZENE) ZENECA.LTD.
 XX
 PI Bird CR, Drake CR, Schuch WW;
 XX
 DR WPI; 1998-042198/04.
 DR P-PSDB; AAW41374.
 XX
 PT Enhancing gene expression without or with reduced co-suppression -
 PT using altered DNA producing different RNA but same protein as
 PT natural gene, useful especially in plants to allow overexpression of
 PT a protein
 XX
 PS Claim 9; Page 15-16; 32pp; English.
 XX
 CC This sequence represents the modified phytoene synthase gene MTOM5, which
 CC is also a chloroplast targeting sequence. This sequence is used in the
 CC method of the invention for enhancing expression of a protein by an
 CC organism, comprising inserting into its genome a nucleotide sequence
 CC which produces different RNA on transcription to that of the gene already

CC present, but produces the same protein on translation. Transgenic plants
 CC with enhanced ability to express a selected can be produced by the
 CC method. For example, the method can be used to achieve overexpression of
 CC a gene specifying an enzyme necessary for carotenoid synthesis in plants
 CC (especially phytoene synthase), to enhance carotenoid expression,
 CC e.g. overexpression of the carotenoid lycopene responsible for the red
 CC colouration of developing tomato fruit. Protein expression is enhanced by
 CC inserting a gene construct which is altered by maximizing the
 CC dissimilarity of nucleotide usage whilst maintaining identity of the
 CC encoded protein. Known methods of increasing protein production by gene
 CC insertion sometimes result in low or no expression (co-suppression),
 CC especially when the recombinant and endogenous gene sequences are
 CC similar. The method allows enhanced expression whilst avoiding or
 CC reducing co-suppression, since sequence similarity between the two genes
 CC is sufficiently reduced.

SQ Sequence 1239 BP; 343 A; 247 C; 320 G; 329 T; 0 other;

Alignment Scores:

Pred. NO.: 1.72e-177 Length: 1239
 Score: 1744.00 Matches: 343
 Percent Similarity: 86.61% Conservative: 32
 Best Local Similarity: 79.21% Mismatches: 34
 Query Match: 76.83% Indels: 24
 DB: 19 Gaps: 3

US-09-847-081b-2 (1-440) x AAV17247 (1-1239)

QY 3 MetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGlyThrGly 22
 DB 1 ATGCGGTGGCAGCTCTTTGGTGGTGAGCCCA---TTCGATGTGTAGTAACGGCACTTCA 57
 QY 23 LeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeuAlaArg 42
 DB 58 TTTATGGAGAGTGTGAGAGAGGTAATAGATTCCTTCGACAGTTCT-----CGT 105
 QY 43 AspArgAsnLeuMetTrpAsnGlyArgLysLysGlyGlyArgGlnArgTrpAsnPhe 62
 DB 106 CACCCTAACCTTGTGTAGTAACGAGCGTATAAACAGGGGA----- 144
 QY 63 GlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGluLysGly 82
 DB 145 -----GGAGGTAACAGCAACAACACGGT 168
 QY 83 SerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGluMetThrValSer 102
 DB 169 AGAAGTTCTCAGTTAGATCAGCAATCCTTGCAACACCTAGCGGTGAGAGAACTATGACT 228
 QY 103 SerGluLysLysValTyrAspValValLeuLysGlnAlaLeuValLysArgGlnLeu 122
 DB 229 AGCGAGCAAAATGGTGTACGACGCTGCTACTTCGTCGAAGCTGCACCTAGTTAAACGTCAGTTA 288
 QY 123 ArgSerThrAspAspLeuGluValLysProAspIleValValProGlyAsnLeuGlyLeu 142
 DB 289 CGTAGTACTACGAACCTTGAGGTTAAACCTTGACATTCCTCAATACCTGGAAACCTTGACAT 348
 QY 143 LeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaLysThrPheTyr 162
 DB 349 CTTTCTGAGGCTTAGCAGACAGATCGCGAGAGGTTTTCGCGAGAAATACCTAAAACCTTCAAT 408
 QY 163 LeuGlyThrLysLeuMetThrProGluArgArgAlaIleTrpAlaIleTyrValTrp 182
 DB 409 TTGGGTACCATGTTGATGACACACAGAAAGCGCTGCTGCAATATATGGCTATTTACGTTTG 468
 QY 183 CysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThrProGlnAla 202
 DB 469 TGTAGGCGTACTGACGAGTTAGTGACGCGACCTAATGCTAGTTACATACACCCGCTGCT 528
 QY 203 LeuAspArgTyrGluThrArgLeuGluAspIlePheSerGlyArgProPheAspMetLeu 222
 DB 529 CTTGACAGATGGAGAACCGTTTGGAGGACGTTTAAACGGCAGACCTTTTCGATATGTTG 588
 QY 223 AspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGlnProPheArgAsp 242

Db	589	GA	CGGAGC	ACTTAGTGTGACACCTGTGAGCAATTTCCCTGTGGACATCCCAACCTTTTCGGGAC	648
QY	243	Met	Ile	leuGluGlyMetArgMetAspLeuThrIpsLysSerArgTyrLysThrPheAspGluLeu	262
Db	649	AT	GATCGAGG	CATGAGATCGATCTTCGTAGTCTCGTTATAGAAATTTTGATGAGTTG	708
QY	263	Tyr	Leu	TyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGly	282
Db	709	TAT	TTTGTACT	GTCTACTGTGCGAGAACCGTGGCGCTTATGTTCAGTCCCTATCATCGGA	768
QY	283	Ile	Ala	ProGluSerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeuGly	302
Db	769	ATT	GCCACGAG	AGGTAAAGCTACTGTAATCTGTTACCCGACGACATAGCAATAGGT	828
QY	303	Leu	Ala	AsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGlyArg	322
Db	829	AT	AGCTAA	CCAGCTACAAATATCTTGAGGACGTGGGTGAGGACGACGTAGGGTCTGT	888
QY	323	Val	Tyr	LeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAla	342
Db	889	GTT	GATCTC	CCACAGACGAGCTCGCTCAAGCTGGATTGAGTGACGAGGACATTTTCGCA	948
QY	343	Gly	Arg	ValThrAspLysTyrPArgAsnPheMetLysLysGlnIleGlnArgAlaArgLys	362
Db	949	GGT	CGTGT	TACAGACAAGTGGAGGATTTTCATGAAAAGCAGATTCACCGTGTCTGTA	1008
QY	363	Phe	Phe	AspGluSerCylLysGlyValThrGluLeuAspSerAlaSerArgTyrProVal	382
Db	1009	TTT	TCGAG	AGAGCTGAAAGGAGTACTGAGCTTCTAGTGCATCAAGTTTCCAGTTTCCAGTT	1068
QY	383	Leu	Thr	AlaLeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAspTyrAsn	402
Db	1069	TGG	CCGAC	CTGTGTCTCTATAGAAGATTTTGGACGAATTCGAGGCTACGATTATAAT	1128
QY	403	Asn	Phe	ThrArgArgAlaTyrValSerLysProLysLysLeuLeuThrLeuProIleAla	422
Db	1129	AAT	TTTAC	TAAAGCTGTACGTTCTTAAGACGAAAAAATATTCGTCTCTTCCAATCGCT	1188
QY	423	Tyr	Ala	LysSerLeuValProProAsnArgThrSerSer 435	
Db	1189	TAC	GCTAA	GAGCTTGGTCCACCAACTAAGACAGCTAGC 1227	
RESULT	9				
AA	99482				
ID	RAZ99482	standard;	cdna;	1239 BP.	
XX	XX				
AC	AAZ99482;				
XX	XX				
DT	03-JUL-2000	(first entry)			
XX	XX				
DE	DE				
XX	XX				
KW	KW	Gibberellic acid; copalyl diphosphate synthase; 3beta-hydroxylase;			
KW	KW	2-oxidase; phytoene synthase; C-20 oxidase; 2beta,3beta-hydroxylase;			
KW	KW	seed germination; seedling growth; gibberellin biosynthetic pathway;			
KW	KW	transgenic plant; hypocotyl; epicotyl; ss.			
XX	XX				
OS	OS	Lycopersicon esculentum.			
XX	XX				
FH	FH	Key	Location/Qualifiers		
FT	FT	1..1239			
FT	FT	/tdg= a			
FT	FT	/product= "phytoene synthase"			
FT	FT	/transl_except= (pos: 1027..1029, aa: Lys)			
FT	FT	/transl_except= (pos: 1057..1059, aa: Arg)			
XX	XX				
PN	PN	WO200005722-A2.			
XX	XX				
PD	PD	24-FEB-2000.			
XX	XX				
PF	PF	10-AUG-1999;	99WO-US18066.		
XX	XX				


```

Db 850 TCATTCTACTGGGACCACCGCTATGACACCGAGAGCGCTTACGTATCTGGCGGATA 909
Qy 180 TyrValTrpCysArgThrAspGluLeuValAspGlyProAsnAlaSerHisIleThr 199
Db 910 TATGATGCTGTAGGAGGACAGATGAGCTTGTGATGGCCCTAACGCGTCACACATAAT 969
Qy 200 ProGlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgProPhe 219
Db 970 CCAACCGCGTGTAGATAGTGGGAACCAAGATTAGAAGATGTTTCAAGGGCAACCTTT 1029
Qy 220 AspMetLeuAspAlaLeuSerAspThrValSerArgPheProValAspIleGlnPro 239
Db 1030 GATAGCTGTGATGCTTATCTGATACCAATACCAAGTATCTGTGGACATCCAGCCA 1089
Qy 240 PheArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTrpLysThrPhe 259
Db 1090 TTTACAGATATAGTAGAAGGAATCGGATGATCTGAAGAATCGAGATACAGATTTTC 1149
Qy 260 AspGluLeuTrpCysTrpTyrValAlaGlyThrValGlyLeuMetSerValPro 279
Db 1150 GATGACGTGATCTTACTGCTATATGCTGCTGCTACAGTTGGCTTGCATGAGTGATCCA 1209
Qy 280 ValMetGlyIleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaLeu 299
Db 1210 GTAATGGGCATGCACTACCTGATCTAAGGCAACAACAGAAAGTGTATATGCACTTTA 1269
Qy 300 AlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAlaArg 319
Db 1270 TCTTTGGGATCGGACACGCTGACTAACATCTTAAGGATGTGGAGAGATGCAAGA 1329
Qy 320 ArgGlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAsp 339
Db 1330 AGAGGAGAGTGTACCTACCTCAAGATGAATAGCACAGCAGGTTTATCAGATGAGGAC 1389
Qy 340 IlePheAlaGlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleArg 359
Db 1390 ATTTTGTGGAAGATGTACAGACAAATGGAGGATTTTATGAAGAACCAATCAAAAGG 1449
Qy 360 AlaArgLysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSerArg 379
Db 1450 GCTAGAAATCTATGATGATGCAGAGAAAGGAGTCCCGCACTCAGCTCCGCGAAGCAG 1509
Qy 380 TrpPro---ValLeuThrAlaLeuLeuTyrArgLysIleLeuAspGluIleGluAla 398
Db 1510 ATTGCTGTGTGGCGCAGCGTTGCTTTTATAGAAAAATATTGATGATGATAGAGA 1569
Qy 399 AsnAspTyrAsnAsnPheThrArgArgAlaTyrValSerLysProLysLysLeuThr 418
Db 1570 AATGACTACACAAATTTACAAAGGGGCTTATGTAAACAAAGGCAAGAAAGCTATTAGCT 1629
Qy 419 LeuProIleAlaTyrAlaLysSerLeu 427
Db 1630 ATGCGTGTAGCATGTGCCAAGTCTCTC 1656

```

RESULT 12

AAV16949
ID AAV16949 standard; cDNA to mRNA; 1921 BP.

XX AC AAV16949;

DT 06-JUL-1998 (first entry)

DE Nucleic acid encoding phytoene synthase 2.

XX Phytoene synthase; breeding; variable flower colour; ds.

OS Gentiana lutea.

XX Key Location/Qualifiers

FT 412..1689

FT CDS

XX /*tag= a

JPL0084966-A.
XX 07-APR-1998.
XX 17-SEP-1996; 96JP-0245107.
XX 17-SEP-1996; 96JP-0245107.
XX (IWAT-) IWATE KEN.
XX WPI: 1998-264853/24.
XX P-PSDB; AAW46962.
XX Phytoene synthase gene - useful for breeding plant of variable flower colour
XX Claim 2; Pages 7-9; 15pp; Japanese.
XX The present sequence encodes phytoene synthase 2. It was isolated from a cDNA library prepared from mRNA extracted from the petals of *Gentiana lutea*. The nucleic acid sequence was amplified from the library using PCR primers AAV16952-53. The phytoene synthase gene is useful for breeding plants with variable flower colours.
XX Sequence 1921 BP; 633 A; 296 C; 438 G; 554 T; 0 other;
Alignment Scores:
Pred. No.: 4,11e-157 Length: 1921
Score: 1557.00 Matches: 305
Percent Similarity: 81.82% Conservatives: 46
Best Local Similarity: 71.10% Mismatches: 68
Query Match: 68.59% Indels: 10
DB: 19 Gaps: 2
US-09-847-081b-2 (1-440) x AAV16949 (1-1921)

Qy 1 MetSerMetSerValAlaLeuLeuTrpValValSerProThrSerGluValSerAsnGly 20
Db 406 GTTAACATGCTATTTGTACGCTATGTTTCCGCGAGTTCTCAAGTTTGTAGTGGC 465
Qy 21 ThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu 40
Db 466 AATGTTTCTTGGAGCCAAATTCGAGAAAGT-----TACCAT 501
Qy 41 AlaArgAspArgAsnLeuMetTrpAsnGlyArgIleLysLysGlyGlyArgGlnArgTrp 60
Db 502 TTTTCGGATAAAAGTTTAAATGTACAAATGGAAGAGTTAAGAAAGTAGACACCAAGCGCT 561
Qy 61 AsnPheGlySerLeuIleAlaAspProArgTyrSerCysLeu-----GlyGlySerArg 78
Db 562 AGATCACGTTATGGGTTGGAGATTGTAGTTTCATTTTGTCTTGAGAGAGTCTGATTAGAG 621
Qy 79 ThrGluLysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyGlu 98
Db 622 ACCCGGGGAAGAAGATTATCGGTATCTCCAGTATTATAGTACCCCGGCGAGAGA 681
Qy 99 MetThrValSerSerGluLysLysValTyrAspValValLeuLysGlnAlaLeuVal 118
Db 682 ATGACGATGACATCAGACAAAGGTTTATGATGTCGTTTAAAGCAACGAGCTTTGATT 741
Qy 119 LysArgGlnLeuArgSerThrAspSerLeuGluValLysProAspIleValProGly 138
Db 742 AATAGACAGTTGAGGCTAGAGAAATTTGGAGGTTGAAACCGGACATTTATTTGCCAGGA 801
Qy 139 AsnLeuGlyLeuLeuSerGluAlaTyrAspArgCysGlyGlyValCysAlaGluTyrAla 158
Db 802 AACCGGAACGTTGAATGAAGCTTATGATCGGTGCGAGAGATGTCGATGATATGCC 861
Qy 159 LysThrPheTyrLeuGlyThrLysLeuMetThrProGluArgArgArgAlaIleTrpAla 178
Db 862 AAGTCATTCCTACTGGGNAACGAGCTCATGACCGGAGAGCGGTTTAGCTATCTGGCG 921
Qy 179 IleTyrValTrpCysArgArgThrAspGluLeuValAspGlyProAsnAlaSerHisIle 198

```
Db 922 ATATATCTATGGTGTAGGACGACAGATGCTGTGTGATGGGCTAAACGGCTCAACATA 981
QY 199 ThrProGlnAlaLeuAspArgTrrpGluThrArgLeuGluAspIlePheSerGlyArgPro 218
Db 982 AATCCACCGCGTTAGATAGTGGGAAGCAAGATTAGAAGATCTTTCAAGGGCAACCT 1041
QY 219 PheAspMetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGln 238
Db 1042 TTTGATATGCTTGTATGCTGTTATCTGATACCATTAACCAAGATATCTGTGGACATCCAG 1101
QY 239 ProPheArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyrLysThr 258
Db 1102 CCATTTAGACATATGATAGGAAGATCGCGATGCTGAGAAATCGAGATACAGAAT 1161
QY 259 PheAspGluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerVal 278
Db 1162 TTCGATGAGCTGATCTTTACTGCTATTATGCTGGTGGTACAGTTGGCTTGATGAGTGA 1221
QY 279 ProValMetGlyIleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAla 298
Db 1222 CCAGTAATGGCATTCACCTGATCTAAGGCAACACAGAAAGTGTGTATATGCGAGCT 1281
QY 299 LeuAlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAla 318
Db 1282 TTATCTTTGGGATCGCGAACACAGCTGACTACATTTAAGGATGTTGGAGAAGATGCA 1341
QY 319 ArgArgGlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGlu 338
Db 1342 AGAAGAGGAAGAGTGTACCTACCTCAAGATGAATTAGCAACAGCAGGTTTATCAGATGAG 1401
QY 339 AspIlePheAlaGlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGln 358
Db 1402 GACATTTTCTGGAAAGATTACAGACAAATGGAGGATTTTATGAAGAACCAATCAAA 1461
QY 359 ArgAlaArgLysPhePheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSer 378
Db 1462 AGGCTAGAAATCTATGATGATGATGAGAAAAGTCCCGGACTCAGCTCCGCGAGC 1521
QY 379 ArgTrpProValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAla 398
Db 1522 AGATTGCTGTGTGGGCGAGCTTGTCTTTTATAGAAAATATTGGATGATAGAAGCA 1581
QY 399 AsnAspTyrAsnAsnAsnPheThrArgArgAlaTyrValSerLysProLysLysLeuThr 418
Db 1582 AATGACTACAACTTTCACAAAAGGGCTTATGTAAACAAAGCAAGCAAGCTATTAGCT 1641
QY 419 LeuProIleAlaTyrAlaLysSerLeu 427
Db 1642 ATGCTGTAGCATGTGCCAAGTCTC 1668

RESULT 13
AAC48162
ID AAC48162 standard; DNA; 1566 BP.
AC AAC48162;
XX
XX
XX 18-OCT-2000 (first entry)
DT
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 56485.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX
```

PR	04-OCT-1999;	99US-01571117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	22-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Alignment Scores:		
Pred. No.:	1.25e-155	Length: 1566
Score:	1542.00	Matches: 318
Percent Similarity:	81.86%	Conservative: 43
Best Local Similarity:	72.11%	Mismatches: 51
Query Match:	67.93%	Indels: 30
DB:	21	Gaps: 8
US-09-847-081B-2 (1-440) x AAC48162 (1-1566)		
Qy	1 MetSerMetSerValAlaLeuLeuTrpValVal-----SerProThrSerGluValSer 18	
Db	153 ATGTCTTCTCTCTGTAGCAGTGTTATGGTGTACTTCTTCTTAATCCAGACCCAATG 212	
Qy	19 AsnGlyThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArg 38	
Db	213 ACAAATTGGGTGGTGA-----AGGTCCTAGAATCTTCTAGA 251	
Qy	39 PheLeuAlaArgAspArgAsnLeuMetTrpAsnGlyArgIleLysLysGly-GlyArgGl 58	
Db	252 CTGTTCTCTCTTGCAG-----AATCAGAGACTAACAAAGSTAAGAAGAG 299	
Qy	58 nArgTrpAsnPheGlySerLeulleAlaAspProArgTyrSerCysLeuGlyGlySerAr 78	
Db	300 CAGATACCAACTTGGAGTTCTTCTT-----TTGTAAGGAACCGA 338	
Qy	78 gThrGluLysGlySerThrPheSerValGlnSerSerLeuValAlaSerProAlaGlyL 98	
Db	339 AGTAGAAGAAAT--GGTGT-----GTGCTTCAGCTTAGTAGCAAGTCCCTTCGGAGA 391	
Qy	98 uMetThrValSerSerGluLysLysValTyrrAspValValLeuLysGlnAlaAlaLeuVa 118	
Db	392 GATAGCTCTTTCATCTGAAGAGAGGTTTACAATGTTGTGTGAACACAGCTGCTTTGT 451	
Qy	118 lLysArgGlnLeuArgSerThrAsp---AspLeuGluValLys-----ProAspIleVa 135	
Db	452 GAACAAACAGCTAAGGTCTTCTTCTTATCACCTTGATGTGAAGAAACCAACAGATCTGT 511	

QY 135 lvalProGlyAsnLeuGlyLeuLeuSerGluAlaTyrAspArgCysGlyGluValCysAl 155
 Db 512 TCCTCTGGAGTTGAGTTTGTGGGTGAAGCTTATGATCATCGGTGAAGTTGGCG 571
 QY 155 aGluTyrAlaLysThrPheTyrLeuGlyThrLysLeuMetThrProGluArgArgAl 175
 Db 572 TGAATATGCTAAGACGCTTTATCTTGAACCTTGTCTTATGACACCGAAGCGAAGGC 631
 QY 175 alleTTPAlaIleTyrValTyrCysArgArgThrAspGluLeuValAspGlyProAsnAl 195
 Db 632 GATTTGGGCAATCTAGCTTGTGTGAAGAAGTGTGAACCTGTGGATGGGCCAAATGC 691
 QY 195 aserHisIleThrProGluAlaLeuAspArgTyrGluThrArgLeuGluAspIlePheSe 215
 Db 692 TTCACATATACTCCATGCTTTAGATAGATGGGAGCAAGGTTAGAAGATCTTTTCGG 751
 QY 215 rGlyArgProPheAspMetLeuAspAlaLeuSerAspThrValSerArgPheProVa 235
 Db 752 TGTCTGCTCTTTCGATATGCTTGTCTGCTGCTGATACAGTTGCTAGATACCCGCT 811
 QY 235 lAspIleGlnProPheArgAspMetIleGluGlyMetArgMetAspLeuThrLysSerAr 255
 Db 812 CGATATTCAGCCATTCGAGACATGATCGAAGGAATGAGATGGACTTGAAGAAATCGAG 871
 QY 255 gTyrLysThrPheAspGluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLe 275
 Db 872 ATACCGAAGCTTCGATGATCTATACCTTTACTGCTACTGCTGCGTGGAACTCGGAT 931
 QY 275 uMetSerValProValMetGlyIleAlaProGluSerLysAlaThrGluSerValTy 295
 Db 932 GATGAGCGTTCGGTGTATGGATCGATCTTAAGTCAAGCAACCAACCAAGTCTTTA 991
 QY 295 rAsnAlaAlaLeuAlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGl 315
 Db 992 CAACGCTGCGCTTGGCTTGGTATGCCAATCACTACTTAACATCTCAGAGACGTAGG 1051
 QY 315 YGluAspAlaArgArgGlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLe 335
 Db 1052 CGAAGATCGAGAGAGAGAGGTTTATCTGCTCAGGATGATTTGGCTCAGGCTGCTCT 1111
 QY 335 uSerAspGluAspIlePheAlaGlyArgValThrAspLysTyrPArgAsnPheMetLysLy 355
 Db 1112 TTCAGATGAGACATATTCGCCGAAAGTAAGTAAATGAGAACTTCATGAAAT 1171
 QY 355 sGlnIleGlnArgAlaArgLysPheAspGluSerGluLysGlyValThrLysLeuAs 375
 Db 1172 GCAGCTTAACGAGCAAGATGTTCTTCGACGAAGCTGAGAAAGGCGTCACCGAGCTCAG 1231
 QY 375 pSerAlaSerArgTyrProValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGl 395
 Db 1232 TGCGCTAGCAGATGCGCTGATGGCTTCATTGCTATTGTACAGAGAACTACTGGACGA 1291
 QY 395 uIleGluAlaAsnAspTyrAsnAsnPheThrArgArgAlaTyrValSerLysProLysLy 415
 Db 1292 GATTGAGCGCAATGATTACAACTTTTACTAAGAGAGCTTATGTGGGAAAGTCAAGAA 1351
 QY 415 sLeuLeuThrLeuProIleAlaTyrAlaLysSerLeuValProProAsnArgThrSerSe 435
 Db 1352 AATGACAGCTTGCCTTATGCTTAATCAGTACTA-----AAGACTTCAAG 1402
 QY 435 r 435
 Db 1403 T 1403
 RESULT 14
 AAC35120
 ID AAC35120 standard; DNA; 1703 BP.
 XX
 AC AAC35120;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 9067.

XX Hybridisation assay; genetic mapping; gene expression control;
 KW Protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 OS Arabidopsis thaliana.
 PN EP1033405-A2.
 XX 06-SEP-2000.
 PD 25-FEB-2000; 2000EP-0301439.
 PF 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.

```
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.

PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 1,41e-155
Score: 1542.00
Percent Similarity: 79.2%
Best Local Similarity: 70.38%
Query Match: 67.93%
DB: 21
Length: 1703
Matches: 316
Conservative: 40
Mismatch: 47
Indels: 46
Gaps: 8

US-09-847-081b-2 (1-440) x AAC35120 (1-1703)

QY 1 MetSerMetSerValAlaLeuLeuTrpValVal-----SerProThrSerGluValSer 18
Db 290 ATGTCTTCTCTGTAGCAGTGTATGGTTGCTACTTCTCTAAATCCAGACCCCAATG 349
QY 19 AsnGlyThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerArg 38
Db 350 AACAAATTGTGGTGGTA-----AGGTTCTAGAAATCTTCTAGA 388
QY 39 PheLeuAlaArgAspArgAsnLeuMetTrpAsnGlyArgIleLysLysGlyGlyArgGln 58
Db 389 CTGTTCTCTCCTGTGCAGAA-----CAG 412
QY 59 ArgTrpAsnPheGlySerLeuLeuAlaAspProArgTrpSerCysLeuGlyGlySerArg 78
Db 413 AGACTAAACAAGGTAAGAGACAGATACCACTTGGAGT----- 454
```

```

QY 79 ThrGluLysGlySerThrPhe-----SerValGlnSer 89
Db 455 -----TCTCTTTTGAATGAACCGAAGTAGAAGATTGGTGTGTGCTTCA 502

QY 90 SerLeuValAlaSerProAlaGlyGluMetThrValSerSerGluLysLysValTyrAsp 109
Db 503 AGCTTAGTACAGTCCCTCTCTGGAGAGATAGCTCTTTCATCTGAAGAGAAGGTTTACAAAT 562

QY 110 ValValLeuLysGlnAlaAlaLeuValLysArgGlnLeuArgSerThrAsp---AspLeu 128
Db 563 GTTGCTTGAACAAGCTCTTGTGGTGAACAACACAGCTAAGTCTCTCTTATGATCTT 622

QY 129 GluValLys-----ProAspLeuValProGlyAsnLeuLysLeuSerGluAla 146
Db 623 GATGTGAAGAACAACCAAGATGTTGTTCTCTCTGGAGTTGAGTTGTTGGTGAAGCT 682

QY 147 TyrAspArgCysGlyGluValCysAlaGluTyrAlaLysThrPheTyrLeuGlyThrLys 166
Db 683 TATGATCGATGCGGTGAAGTTGCGGTGAATATGCTAAGACCTTTTATCTTGAACCTTG 742

QY 167 LeuMetThrProGluArgArgAlaIleTyrAlaIleTyrValTyrCysArgArgThr 186
Db 743 CTTATGACACCCGAAGCGAAGCGATTGGGCTATCTAGCTTGGTGTAGAAGAACT 802

QY 187 AspGluLeuValAspGlyProAsnAlaSerHisIleThrProGlnAlaLeuAspArgTyr 206
Db 803 GATGAACCTTGTGGATGGCCAAATGCTTCACATATAACTCCCATGCTTTAGATAGATGG 862

QY 207 GluThrArgLeuGluAspIlePheSerGlyArgProPheAspMetLeuAspAlaLeu 226
Db 863 GAAGCAAGGTTAGAGATCTTTCCGTGCTGCTCTTCATGATGCTTGCATGCTGCTC 922

QY 227 SerAspThrValSerArgPheProValAspIleGlnProPheArgAspMetIleGluGly 246
Db 923 GCTGATACAGTTGCTAGATACCTGCGATATTCAGCCATTCGAGACATGTCGAAGA 982

QY 247 MetArgMetAspLeuTyrLysSerArgTyrLysThrPheAspGluLeuTyrLysCys 266
Db 983 ATGAGAAATGGACTTGAAGAATCGAGATACCAAGACTTCGATGATCATACCTTACTGC 1042

QY 267 TyrTyrValAlaGlyThrValGlyLeuMetSerValProValMetGlyIleAlaProGlu 286
Db 1043 TACTAGTGTGCGGAACCTCGGATTCGAGCGTTCGCGTTATGGATCGATCCTAAG 1102

QY 287 SerLysAlaThrThrGluSerValTyrAsnAlaAlaLeuAlaLeuGlyLeuAlaAsnGln 306
Db 1103 TCGAAAGCAACACCCGAAAGTGTTCACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1162

QY 307 LeuThrAsnIleLeuArgAspValGlyGluAspAlaArgArgGlyArgValTyrLeuPro 326
Db 1163 CTTACTAATACATCTAGACAGCTGAGCGAAGATGCGAGAAGAGGAGGTTTATCTGCT 1222

QY 327 GlnAspGluLeuAlaGlnAlaGlyLeuSerAspGluAspIlePheAlaGlyArgValThr 346
Db 1223 CAGGATGAATGGCTCAGCGTGGTCTTCAGATGAAGACATATTCGCCGGAAGATTA 1282

QY 347 AspLysTrpArgAsnPheMetLysLysGlnIleGlnArgAlaArgLysPheAspGlu 366
Db 1283 GATAAATGAGAACTTCATGAAATGACAGTAAACAGCAAGAAATGTTCTTCGACGAA 1342

QY 367 SerGluLysGlyValThrGluLeuAspSerAlaSerArgTyrProValLeuThrAlaLeu 386
Db 1343 GCTGAGAAGGCGTCAACGAGTCACTGCGGTAGCAGATGCGCTGATGGGCTTCATTG 1402

QY 387 LeuLeuTyrArgLysIleLeuAspGluIleGluAlaAsnAspTyrAsnAsnPheThrArg 406
Db 1403 CTATTGTACAGGAGATCTGACGAGATGAGCGAAGTGAAGCGAAGTATACCAATTTTACTAAG 1462

QY 407 ArgAlaTyrValSerLysProLysLysLeuLeuThrLeuProIleAlaTyrAlaLysSer 426
Db 1463 AGAGCTTATGTGGGGAAGTCAAGAAATTCGACGCTTTCGCCATTCGCTTATGCTAATCA 1522

```

```

QY 427 LeuValProAsnArgThrSerSer 435
Db 1523 GFACTA-----AAGACTTCAAGT 1540

RESULT 15
AAV16948
ID AAV16948 standard; cDNA to mRNA; 2085 BP.
XX AAV16948;
XX 06-JUL-1998 (first entry)
XX Nucleic acid encoding phytoene synthase 1.
XX Phytoene synthase; breeding; variable flower colour; ds.
XX Gentiana lutea.
XX Key Location/Qualifiers
XX FT 586..1866
XX CDS /*tag= a
XX JP10084966-A.
XX 07-APR-1998.
XX 17-SEP-1996; 96JP-0245107.
XX 17-SEP-1996; 96JP-0245107.
XX (IWAT-) IWATE KEN.
XX WPI; 1998-264853/24.
XX P-PSDB; AAW46961.
XX phytoene synthase gene - useful for breeding plant of variable.
XX flower colour
XX Claim 1; Pages 5-7; 15pp; Japanese.
XX The present sequence encodes phytoene synthase 1. It was isolated from
XX a cDNA library prepared from mRNA extracted from the petals of Gentiana
XX lutea. The nucleic acid sequence was amplified from the library using
XX PCR primers AAV16952-53. The phytoene synthase gene is useful for
XX breeding plants with variable flower colours.
XX SQ Sequence 2085 BP; 654 A; 320 C; 479 G; 632 T; 0 other;

Alignment Scores:
Pred. No.: 2,89e-154 Length: 2085
Score: 1531.00 Matches: 310
Percent Similarity: 82.56% Conservative: 45
Best Local Similarity: 72.09% Mismatches: 64
Query Match: 67.44% Indels: 12
DB: 19 Gaps: 3

US-09-847-081B-2 (1-440) x AAV16948 (1-2085)
QY 1 MetSerMetSerValAlaLeuLeuTyrValValSerProThrSerGluValSerAsnGly 20
Db 580 GTTACATGCTCTATTGTGTAGCGTATGGTGTGTTCCCGAGTTCCTGAAGTTTGTAGTGGC 639
QY 21 ThrGlyLeuLeuAspSerValArgGluGlyAsnArgValPheValSerSerArgPheLeu 40
Db 640 AATGTTTCTTGGAGCAATTGAGAACT-----TACCAT 675
QY 41 AlaArgAspArgAsnLeuMetThrAsnGlyArgIleLysLysGlyGlyArgGlnArgTyr 60
Db 676 TTTTCGGATAAAAGTTTAAATGTACAAATGGAAGATTAAGAAAAGTAGACACCAAGCGT 735
QY 61 AsnPheGlySerLeuIleAlaAspProArgTyrSerCysLeuGlyGlySerArgThrGlu 80
Db 736 AGATCACGTTATGGGGTGGAGATTGAGTTCATTTTGTCTTGAGAGAGTCTGATTAGAG 795

```

Search completed: April 5, 2003, 03:09:28
Job time : 306 secs

```
Qy 81 ---LysGlySerThrPheSer-ValGlnSerSerLeuValAlaSerProAlaGlyGluMe 99
Db 796 ACCCGGGGAAGAAGATTATCGGTATCTCCATATTAGTACCCCGCAGGAGAAAT 855
Qy 99 tThrValSerSerGluLysLysValTyrAspValValLeuLysGlnAlaAlaLeuVally 119
Db 856 GACATG-ACATCAGAGCAAAAGGTTTATGATGTCGTTTAAAGCAACGACGTTTGATTA 914
Qy 119 sArgGlnLeuArgSerThrAspAspLeuGluValLysProAspIleValValProGlyAs 139
Db 915 TAGACAGTTGAGTCTAGAGAAATTTGGAGGTGAACCCGACACATTATTTTGGCAGGAAA 974
Qy 139 nLeuGlyLeuLeuSerGluAlaTyrAspArgCysGlyGluValCysAlaGluTyrAlaLy 159
Db 975 CGCGAACGTTGTAATGAAGCTTATGATCGGTGTCGAGAAATGCTGCTGAATATGCCAA 1034
Qy 159 sThrPheTyrLeuGlyThrLysLeuMetThrProGluArgArgAlaIleTrpAla-- 178
Db 1035 GTCAATCTACTGGGAACCCAGCTCATGACCGGAGGCGGTTTAGCTATCTGGGCCGA 1094
Qy 179 -IleTyrValTrpCysArgThrAspGluLeuValAspGlyProAsnAlaSerHisIi 198
Db 1095 TATATATGTTGGGTAGGAGACAGATGAGTTGTTGATGGGCTTAACCGCTCACACAT 1154
Qy 198 eThrProGlnAlaLeuAspArgTrpGluThrArgLeuGluAspIlePheSerGlyArgPr 218
Db 1155 AAATCCACCCGCGTTAGTAGTGGGAAGCAAGATCAGAAGATGTTTTCACAGGGCAACC 1214
Qy 218 oPheAspMetLeuAspAlaAlaLeuSerAspThrValSerArgPheProValAspIleGl 238
Db 1215 TCTGGATATGCTGATGCTGCTTATCTGATACCATTTACCAAGTATCTCTGGACATCCA 1274
Qy 238 nProPheArgAspMetIleGluGlyMetArgMetAspLeuTrpLysSerArgTyrLysTh 258
Db 1275 GCCATTTAGAGATATGATAGAGGATGCGGATGGATCTGAAGNAATCGAGATACAAGAA 1334
Qy 258 rPheAspGluLeuTyrLeuTyrCysTyrTyrValAlaGlyThrValGlyLeuMetSerVa 278
Db 1335 TTTGATGAGCTGATCTTACTGCTATTATGTGGCTGGCACAGTTGGCTTGATGAGTGT 1394
Qy 278 lProValMetGlyIleAlaProGluSerLysAlaThrThrGluSerValTyrAsnAlaAl 298
Db 1395 ACCAGTAATGGGATTCGACCTGAATCTAAGGCAACACAGAAAGTGTGTATGATGCAGC 1454
Qy 298 aLeuAlaLeuGlyLeuAlaAsnGlnLeuThrAsnIleLeuArgAspValGlyGluAspAl 318
Db 1455 TTTATCTTTGGGATCCGACCAACGACGCTACTACATTTCTAAGGGATGTTGGAGAGATGC 1514
Qy 318 aArgArgGlyArgValTyrLeuProGlnAspGluLeuAlaGlnAlaGlyLeuSerAspGl 338
Db 1515 AAGAAGAGGGAGAGTGATCCTACTCAAGATGAATTAGCACAAAGCAGGTTTATCAGATGA 1574
Qy 338 uAspIlePheAlaGlyArgValThrAspLysTrpArgAsnPheMetLysLysGlnIleGl 358
Db 1575 GGACATTTTGGCGAAAGATTACAGACAAATGGAGATTTTTATGAAGAAGCAAAATCAA 1634
Qy 358 nArgAlaArgLysPheAspGluSerGluLysGlyValThrGluLeuAspSerAlaSe 378
Db 1635 AAGGCTAGAAAATCTATGATGATGACGAGAAAGGAGTCCCGAAGTCTGCTCCGCGAG 1694
Qy 378 rArgTrpProValLeuThrAlaLeuLeuLeuTyrArgLysIleLeuAspGluIleGluAl 398
Db 1695 CAGATTGCTGTGTGGCAGCGTTGCTTTTATAGAAAAATATTTGGATCAGATAGAGC 1754
Qy 398 aAsnAspTyrAsnAsnPheThrArgArgAlaTyrValSerLysProLysLysLeuLeuTh 418
Db 1755 AAATGACTACAAATTTCCCGCAGAGGGCTTATGTAAACGAGCGGAAGAGCTATTAGC 1814
Qy 418 rLeuProIleAlaTyrAlaLysSerLeu 427
Db 1815 TATGCCCTGTAGCTGTGCCAAGTCTCTC 1842
```